Tight Information-Theoretic Lower Bounds for Welfare Maximization in Combinatorial Auctions

Vahab Mirrokni [*]   Michael Schapira [†]   Jan Vondrák[‡]

November 14, 2007

Abstract

We provide tight information-theoretic lower bounds for the welfare maximization problem in combinatorial auctions. In this problem the goal is to partition m items between k bidders in a way that maximizes the sum of bidders' values for their allocated items. Bidders have complex preferences over items expressed by valuation functions that assign values to all subsets of items.

We study the "black box" setting in which the auctioneer has oracle access to the valuation functions of the bidders. In particular, we explore the well-known value queries model in which the permitted query to a valuation function is in the form of a subset of items, and the reply is the value assigned to that subset of items by the valuation function.

We consider different classes of valuation functions: Submodular, subadditive, and superadditive. For these classes it has been shown that one can achieve approximation ratios of $1 - \frac{1}{e}$, $\frac{1}{\sqrt{m}}$, and $\frac{\sqrt{\log m}}{m}$, respectively, via a polynomial (in n and m) number of value queries. We prove that these approximation factors are essentially the best possible: For any fixed $\epsilon > 0$, a $(1 - 1/e + \epsilon)$-approximation for submodular valuations or an $\frac{1}{m^{1/2-\epsilon}}$-approximation for subadditive valuations would require exponentially many value queries, and a $\frac{\log^{1+\epsilon} m}{m}$-approximation for superadditive valuations would require a superpolynomial number of value queries.

[*]Theory Group, Microsoft Research, Redmond, WA 98052, E-mail: `mirrokni@microsoft.com`

[†]The School of Computer Science and Engineering, The Hebrew University of Jerusalem, Israel. E-mail:`mikesch@cs.huji.ac.il`. Supported by grants from the Israel Science Foundation and the USA-Israel Binational Science Foundation.

[‡]Dept. of Mathematics, Princeton, NJ 08544, USA. E-mail: `jvondrak@math.princeton.edu`

1 Introduction

Combinatorial auctions are a central research area at the intersection of in economics, game theory, and computer science. The welfare maximization problem in combinatorial auctions is an abstraction of many computational and economic resource-allocation problems. In this problem, an auctioneer sells a set M of m items to a set N of k bidders. The value of bidder $i \in N$ for any subset (bundle) of items is given by a valuation set function $v_i : 2^M \to \mathbb{R}_+$, where $v_i(S)$ represents i's maximum willingness to pay for the bundle S. The two standard assumptions on each v_i are that if $S \subseteq T$ then $v_i(S) \leq v_i(T)$ (monotonicity), and that $v_i(\emptyset) = 0$ (normalization). The objective is to partition M into k disjoint subsets $S_1, S_2, \ldots, S_k$ in a way that maximizes the expression $\sum_{1 \leq i \leq k} v_i(S_i)$ (i.e., the social welfare).

Algorithms for maximizing welfare in combinatorial auctions are required to be polynomial in the natural parameters of the problem, m and k. However, since the "input" (the valuation functions) is of exponential size one must specify how it can be accessed. Most works in this field have taken a "black box" approach in which bidders' valuation functions are accessed via oracles that can answer specific type of queries. Three types of queries have been considered [1, 2, 5]:

- **Value queries:** The query to a valuation function v_i is in the form of a bundle $S \subseteq M$, and the response is $v_i(S)$.

- **Demand queries:** The query to a valuation function v_i is in the form of a price vector $p = (p_1, ..., p_m)$ and the response is the bundle T most demanded by v_i given these prices. That is, $T = \mathrm{argmax}_{S \subseteq M} v_i(S) - \sum_{j \in S} p_j$.

- **General queries:** We allow any type of query (to single valuation functions). This model captures the communication complexity (between the bidders) of the problem. Due to its strength it is mostly interesting for proving lower bounds.

Value queries are strictly less powerful than demand queries, which, in turn, are strictly less powerful than general queries [1, 2, 14]. In this paper we focus our attention on the value-queries model.

It has been shown that computing an optimal solution for social-welfare maximization requires an exponential number of queries even in the general queries model [13]. Hence, we are interested in the approximability of this problem. For the general case, the approximability of the problem is well-understood in all models [1, 13]. Researchers have also studied the approximability of the problem for restricted classes of valuation functions. Two families of such classes, that have natural economic interpretations [11, 5], have been considered: *Subadditive* functions, and *superadditive* functions. A set function $f : 2^M \to \mathbb{R}_+$ is subadditive iff for any two sets S and T, $f(S) + f(T) \geq f(S \cup T)$, and is superadditive iff for any two *disjoint* sets S and T, $f(S) + f(T) \leq f(S \cup T)$. An important special case of subadditive functions are *submodular* functions. A set function f is submodular iff for any two sets S and T, $f(S) + f(T) \geq f(S \cup T) + f(S \cap T)$. Submodular functions (that are monotone and normalized) are also reasonable to consider from an economic perspective as they characterize functions with decreasing marginal utilities [11].

We present tight information-theoretic lower bounds for submodular, subadditive, and super-additive valuation functions in the value query model. We prove the following theorems:

Theorem: For any fixed $\epsilon > 0$, achieving an approximation ratio of $1 - \frac{1}{e} + \epsilon$ for welfare-maximization with submodular functions requires an exponential number of value queries.

This matches the $(1 - \frac{1}{e})$-approximation (achieved with a polynomial number of value queries) recently shown by Vondrák [17], who improved over the $\frac{1}{2}$-approximation shown by Lehmann et

al. [11]. (We note that this problem can be formalized as the problem of maximizing a submodular function subject to matroid constraints. Hence, the greedy algorithm developed by Fisher et al. [16] provides a $\frac{1}{2}$-approximation for this problem). The only previously known information-theoretic lower bound for this problem was $1 - O(\frac{1}{m})$, in the general queries model. Our lower bound strengthens the $1 - \frac{1}{e} + \epsilon$ lower bound, dependent on $P \neq NP$ proven by Khot et al. [10]. We stress that our lower bound is independent of any computational complexity assumptions and holds even for algorithms of unbounded computational power, that are bounded only in terms of the number of value queries they can make. Also, we remark that the same inapproximability result does not hold in stronger query models – it is known that $(1 - 1/e + \epsilon)$-approximation is possible with polynomially many demand queries [7].

Theorem: For any fixed $\epsilon > 0$, achieving an approximation ratio of $\frac{1}{m^{\frac{1}{2}-\epsilon}}$ for welfare-maximization with subadditive functions requires an exponential number of value queries.

This matches the upper bound of $\frac{1}{\sqrt{m}}$ presented by Dobzinski et. al [5] (achieved using a polynomial number of value queries). The previously known lower bound in the value-queries model was $\frac{1}{m^{\frac{1}{4}}}$ [4]. In fact, our lower bound holds even for the more restricted subclass of fractionally subadditive valuations [6], introduced in [12] under the name of XOS.

Theorem: For any fixed $\epsilon > 0$, achieving an approximation ratio of $\frac{\log m^{1+\epsilon}}{m}$ for welfare-maximization with superadditive functions requires a super-polynomial number of value queries.

This nearly matches the upper bound of $\frac{\sqrt{\log m}}{m}$ presented by Holzman et. al [9] (achieved via a polynomial number of value queries). A similar lower bound was known for general valuation functions [1, 2]. We extend this lower bound to the restricted class of superadditive functions. In fact, the lower bound holds for a superadditive analogue of fractionally subadditive functions (that is strictly contained in the class of superadditive valuation functions).

2 Value-Query Complexity of Submodular Welfare Maximization

In this section, we construct an example showing that it is impossible to achieve an approximation factor better than $1 - 1/e$ for the submodular welfare problem, using a polynomial number of value queries. More precisely, we prove the following.

Theorem 2.1. *For any fixed $\beta > 0$ and $k \geq 2$, a $(1 - (1 - 1/k)^k + \beta)$-approximation algorithm for k players with submodular valuation functions requires exponentially many value queries.*

Since $1 - (1 - 1/k)^k$ is arbitrarily close to $1 - 1/e$ for large enough k, this implies the following.

Corollary 2.2. *For any fixed $\beta > 0$, there is no $(1 - 1/e + \beta)$-approximation for an arbitrary number of players, using a subexponential number of queries.*

We note that our examples use the same submodular valuation function for all players, just like in the NP-hardness result of [10]. Our construction is different, however. While the hardness reduction of [10] uses explicit coverage-type submodular functions, our valuation functions are not exactly of the coverage type. In fact, our construction is inspired by a lower bound developed by Feige et. al [8] for inapproximability of maximizing non-monotone submodular functions. In the following, we describe the construction.

Consider a k-uniform hypergraph $H = (X, E)$ and a function $f : 2^X \to \mathbb{R}_+$ where $f(S)$ is the number of hyperedges incident with the set of vertices S. This is a coverage-type submodular

function. The idea is that it is hard to distinguish instances where H is a complete k-partite hypergraph (and allocating one part to each player results in a "perfect solution"), and instances where H is a complete k-uniform hypergraph (and then there is no "perfect solution"). But in order to make the example work, we have to modify the coverage-type functions slightly.

We consider a ground set X partitioned into $X_1 \cup X_2 \cup \ldots \cup X_k$. The functions $f(S)$ that we define depend only on the fractions of X_i that S contains: $x_i = |S \cap X_i|/|X_i|$. To simplify the notation, we work with continuous functions $\tilde{f}(x_1, \ldots, x_n)$. The following lemma states the properties that we need $\tilde{f}(x_1, \ldots, x_n)$ to satisfy.

Lemma 2.3. *Let $X = X_1 \cup X_2 \cup \ldots \cup X_n$ as above and let $\tilde{f} : [0,1]^n \to \mathbb{R}$ be a function with continuous first partial derivatives, and second partial derivatives almost everywhere. Define a discrete function $f : X \to \mathbb{R}$ so that*

$$f(S) = \tilde{f}\left(\frac{|S \cap X_1|}{|X_1|}, \ldots, \frac{|S \cap X_n|}{|X_n|}\right).$$

1. If $\frac{\partial \tilde{f}}{\partial x_i} \geq 0$ everywhere for each i, then the function f is monotone.

2. If $\frac{\partial^2 \tilde{f}}{\partial x_i \partial x_j} \leq 0$ almost everywhere [1] for any i, j, then the function f is submodular.

Proof. For monotonicity, it's sufficient to observe that if $\frac{\partial \tilde{f}}{\partial x_i} \geq 0$, then $\tilde{f}$ is non-decreasing in each coordinate. Hence, adding elements cannot decrease the value of f.

For the submodularity condition, fix an element in $a \in X_i$ and consider a set S parameterized by $x_i = |S \cap X_i|/|X_i|$. The marginal value of a added to S is equal to

$$\begin{aligned}
f_S(a) &= \tilde{f}(x_1, \ldots, x_i + \frac{1}{|X_i|}, \ldots, x_n) - \tilde{f}(x_1, \ldots, x_i, \ldots, x_n) \\
&= \int_0^{1/|X_i|} \frac{\partial \tilde{f}}{\partial x_i}(x_1, \ldots, x_i + t, \ldots, x_n) dt.
\end{aligned}$$

We want to prove that $f_S(a)$ cannot increase by adding elements to S, i.e. by increasing any coordinate x_j. Because $\frac{\partial \tilde{f}}{\partial x_i}$ is continuous and its derivative along x_j, $\frac{\partial^2 \tilde{f}}{\partial x_i \partial x_j}$, is non-positive except at finitely many points, $\frac{\partial \tilde{f}}{\partial x_i}$ is non-increasing in x_j. By shifting the entire integral to a higher value of x_j, the marginal value cannot increase. $\square$

Hence, we need our continuous functions to satisfy $\frac{\partial f}{\partial x_i} \geq 0$ and $\frac{\partial^2 f}{\partial x_i \partial x_j} \leq 0$ for all $i, j \in [k]$, which implies monotonicity and submodularity in the discrete case. We call such functions *smooth submodular*. To shorten notation, we write $f(\mathbf{x}) = f(x_1, \ldots, x_n)$.

In each instance, all players have the same valuation function. We find two functions f, g such that we have $f(\mathbf{x}) = g(\mathbf{x})$ whenever $\max_{i,j} |x_i - x_j| \leq \epsilon$. This will be the case for any subexponential number of queries with high probability, assuming that the underlying partition is random and unknown to the algorithm. Then a typical query is partitioned into k equal parts by $X_1, \ldots, X_k$, and a deviation of constant ϵ from the expectation occurs with exponentially small probability, which can be proved by standard Chernoff bounds. (It should be noted here that we keep k and ϵ constant, while the number of vertices n tends to infinity.) Therefore, f and g are indistinguishable by a subexponential number of queries. If the gap between the optima of f and g is close to $1 - (1 - 1/k)^k$, this implies that any approximation algorithm improving this factor would require an exponential number of value queries. Therefore it remains to prove the following.

---

[1] To be more precise, on any axis-parallel line there are only finitely many points where $\frac{\partial^2 \tilde{f}}{\partial x_i \partial x_j}$ is not defined.

3

Lemma 2.4. *For any $\beta > 0$ and integer $k \geq 2$, there is $\epsilon > 0$ and two smooth submodular functions $f, g : [0, 1]^k \to \mathbb{R}_+$ such that*

- *If $\max_{i,j} |x_i - x_j| \leq \epsilon$, then $f(\mathbf{x}) = g(\mathbf{x})$.*

- $\max\{\sum_i f(x_{i1}, \ldots, x_{ik}) \mid \forall j; \sum_i x_{ij} = 1\} \geq (1 - \beta)k.$

- $\max\{\sum_i g(x_{i1}, \ldots, x_{ik}) \mid \forall j; \sum_i x_{ij} = 1\} \leq (1 - (1 - 1/k)^k + \beta)k.$

Proof. We start by considering two smooth submodular functions, motivated by the examples of k-uniform hypergraphs that we discussed above.

- $f(\mathbf{x}) = 1 - \prod_{i=1}^{k}(1 - x_i).$

- $g(\mathbf{x}) = 1 - (1 - \bar{x})^k$, where $\bar{x} = \frac{1}{k}\sum_{i=1}^{k} x_i.$

The optimal solution with valuation function f is $x_{ii} = 1$, $x_{ij} = 0$ for $i \neq j$. This way, each player gets the maximum possible value 1. For g, on the other hand, the value depends only on the average of the coordinates $\bar{x}$. By the concavity of g, the optimum solution is to allocate $\bar{x} = 1/k$ to each player, which gives her a value of $1 - (1 - 1/k)^k$.

It remains to perturb the functions so that $f(\mathbf{x}) = g(\mathbf{x})$ for vectors satisfying $\max_{i,j} |x_i - x_j| \leq \epsilon$. Let $h(\mathbf{x})$ denote the difference of the two functions,

- $h(\mathbf{x}) = f(\mathbf{x}) - g(\mathbf{x}) = (1 - \bar{x})^k - \prod_{i=1}^{k}(1 - x_i).$

Again, we denote $\bar{x} = \frac{1}{k}\sum_{i=1}^{k} x_i$. Also, let $\delta = \max_{i,j} |x_i - x_j|$. First, we estimate $h(\mathbf{x})$ and its first derivatives in terms of $\bar{x}$ and δ. We use very crude bounds, to simplify the analysis.

Claim.

1. $h(\mathbf{x}) \leq k\delta(1 - \bar{x})^{k-1}.$

2. $h(\mathbf{x}) \geq k^{-4}\delta^2(1 - \bar{x})^{k-2}.$

3. $|\frac{\partial h}{\partial x_j}| \leq k\delta(1 - \bar{x})^{k-2}$, i.e. $|\frac{\partial h}{\partial x_j}| \leq k^3(1 - \bar{x})^{k/2-1}\sqrt{h(\mathbf{x})}.$

1. We have $h(\mathbf{x}) = (1 - \bar{x})^k - \prod_{i=1}^{k}(1 - x_i)$. If $k\delta \geq 1 - \bar{x}$, we get immediately $h(\mathbf{x}) \leq (1 - \bar{x})^k \leq k\delta(1 - \bar{x})^{k-1}$. So let's assume $k\delta < 1 - \bar{x}$. Then, since $x_i \leq \bar{x} + \delta$ for all i, we get

$$h(\mathbf{x}) \leq (1 - \bar{x})^k - (1 - \bar{x} - \delta)^k = (1 - \bar{x})^k \left(1 - \left(1 - \frac{\delta}{1 - \bar{x}}\right)^k\right) \leq (1 - \bar{x})^k \frac{k\delta}{1 - \bar{x}}.$$

2. For a lower bound on $h(\mathbf{x})$, suppose that $\delta = x_2 - x_1$ and define $\eta = \frac{1}{k-2}(\bar{x} - \frac{1}{2}(x_1 + x_2))$. I.e., $x_1 = \bar{x} - (k - 2)\eta - \delta/2$, $x_2 = \bar{x} - (k - 2)\eta + \delta/2$, and the average of the remaining coordinates is $\bar{x} + 2\eta$. By the arithmetic-geometric mean inequality, $\prod_{i \neq 1,2}(1 - x_i)$ is maximized when these variables are all equal:

$$
\begin{aligned}
h(\mathbf{x}) &\geq (1 - \bar{x})^k - (1 - \bar{x} + (k - 2)\eta + \frac{1}{2}\delta)(1 - \bar{x} + (k - 2)\eta - \frac{1}{2}\delta)(1 - \bar{x} - 2\eta)^{k-2} \\
&= (1 - \bar{x})^k - (1 - \bar{x} + (k - 2)\eta)^2(1 - \bar{x} - 2\eta)^{k-2} + \frac{1}{4}\delta^2(1 - \bar{x} - 2\eta)^{k-2}.
\end{aligned}
$$

4

Again by the arithmetic-geometric inequality, the first term is always larger than the second term. If $\eta \le \frac{1}{k}(1 - \bar{x})$, we are done because then the last term is at least $\frac{1}{4e^2}\delta^2(1 - \bar{x})^{k-2}$. So we can assume $\eta > \frac{1}{k}(1 - \bar{x})$. In this case, we throw away the last term and write

$$
\begin{aligned}
h(\mathbf{x}) &\ge (1 - \bar{x})^k - (1 - \bar{x} + (k - 2)\eta)^2 (1 - \bar{x} - 2\eta)^{k-2} \\
&= (1 - \bar{x})^k \left(1 - \left(1 + (k - 2)\frac{\eta}{1 - \bar{x}}\right)^2 \left(1 - \frac{2\eta}{1 - \bar{x}}\right)^{k-2}\right) \\
&\ge (1 - \bar{x})^k \left(1 - \left(1 + \frac{\eta}{1 - \bar{x}}\right)^{2(k-2)} \left(1 - \frac{\eta}{1 - \bar{x}}\right)^{2(k-2)}\right) \\
&= (1 - \bar{x})^k \left(1 - \left(1 - \frac{\eta^2}{(1 - \bar{x})^2}\right)^{2(k-2)}\right) \\
&\ge (1 - \bar{x})^k \left(1 - \left(1 - \frac{1}{k^2}\right)^{2(k-2)}\right) \ge \frac{1}{k^2}(1 - \bar{x})^k.
\end{aligned}
$$

We observe it always holds that $\delta \le k(1 - \bar{x})$: If the minimum coordinate is x_{min}, we have $\bar{x} \le \frac{1}{k}x_{min} + \frac{k-1}{k} \cdot 1$, hence $x_{min} \ge k\bar{x} - (k - 1)$ and $\delta \le 1 - x_{min} \le k(1 - \bar{x})$.
Consequently, $h(\mathbf{x}) \ge k^{-2}(1 - \bar{x})^k \ge k^{-4}\delta^2(1 - \bar{x})^{k-2}$.

3. Let $\delta = \max_{i,j} |x_i - x_j|$. We estimate the partial derivative

$$
\frac{\partial h}{\partial x_j} = \prod_{i \ne j}(1 - x_i) - (1 - \bar{x})^{k-1}.
$$

Define $\eta = \frac{1}{k-1}(x_j - \bar{x})$. I.e., $x_j = \bar{x} + (k - 1)\eta$ and the average of the remaining coordinates is $\bar{x} - \eta$. By the arithmetic-geometric inequality,

$$
\frac{\partial h}{\partial x_j} \le (1 - \bar{x} + \eta)^{k-1} - (1 - \bar{x})^{k-1} = (1 - \bar{x})^{k-1}\left(\left(1 + \frac{\eta}{1 - \bar{x}}\right)^{k-1} - 1\right).
$$

Since $\eta = \frac{1}{k-1}(x_j - \bar{x}) \le \frac{1}{k-1}(1 - \bar{x})$, we can estimate $(1 + \frac{\eta}{1-\bar{x}})^{k-1} \le 1 + 2k\frac{\eta}{1-\bar{x}}$. Also, we know that all coordinates differ from $\bar{x} - \eta$ by at most δ, in particular $x_j = \bar{x} + (k - 1)\eta \le \bar{x} - \eta + \delta$, hence $k\eta \le \delta$ and

$$
\frac{\partial h}{\partial x_j} \le (1 - \bar{x})^{k-1} \cdot 2k\frac{\eta}{1 - \bar{x}} \le 2\delta(1 - \bar{x})^{k-2}.
$$

For a lower bound, it's enough to observe that each coordinate is at most $\bar{x} + \delta$, and so

$$
\begin{aligned}
\frac{\partial h}{\partial x_j} &\ge (1 - \bar{x} - \delta)^{k-1} - (1 - \bar{x})^{k-1} \\
&= (1 - \bar{x})^{k-1}\left(\left(1 - \frac{\delta}{1 - \bar{x}}\right)^{k-1} - 1\right) \\
&\ge (1 - \bar{x})^{k-1}\left(-(k - 1)\frac{\delta}{1 - \bar{x}}\right) \\
&= -(k - 1)\delta(1 - \bar{x})^{k-2}
\end{aligned}
$$

assuming that $(k - 1)\frac{\delta}{1-\bar{x}} \le 1$; otherwise we get the same bound directly from $\frac{\partial h}{\partial x_j} \ge -(1 - \bar{x})^{k-1}$. This finishes the proof of the claim.

We return to our construction. We define $\tilde{f}(\mathbf{x}) = f(\mathbf{x}) - \phi(h(\mathbf{x}))$ where $\phi : \mathbb{R} \to \mathbb{R}$ is defined so that $\phi(t) = t$ for small $t \geq 0$, then $\phi(t)$ is increasing and concave with a controlled second derivative and finally $\phi(t)$ is bounded by a small constant everywhere. More precisely,

- For $t \in [0, \epsilon_1]$, we set $\phi(t) = t$. We choose $\epsilon_1 = k\epsilon$. I.e., for $\max_{i,j} |x_i - x_j| \leq \epsilon$, we have $h(\mathbf{x}) \leq \epsilon_1$ by Claim 1 and then $\tilde{f}(\mathbf{x}) = g(\mathbf{x})$.

- For $t \in [\epsilon_1, \epsilon_2]$, the first derivative of ϕ is continuous at $t = \epsilon_1$ and its second derivative is $\phi''(t) = -\alpha/t$ for $t \in [\epsilon_1, \epsilon_2]$. Hence,

$$\phi'(t) = 1 - \int_{\epsilon_1}^{t} \frac{\alpha}{\tau} d\tau = 1 - \alpha \ln \frac{t}{\epsilon_1}.$$

We choose $\alpha = 2/\ln \frac{1}{\epsilon_1}$ and $\epsilon_2 = \sqrt{\epsilon_1}$, so that $\phi'(\epsilon_2) = 0$. Since $\phi'(t) \leq 1$ everywhere, we have $\phi(\epsilon_2) \leq \epsilon_2$.

- For $t > \epsilon_2$, we set $\phi(t) = \phi(\epsilon_2)$.

Hence, we have $0 \leq \phi(t) \leq \epsilon_2$ everywhere and $\tilde{f}(\mathbf{x}) = f(\mathbf{x}) - \phi(h(\mathbf{x})) \geq f(\mathbf{x}) - \epsilon_2$. Next, we want to show that we didn't corrupt the monotonicity and submodularity of f too badly. We have

$$\frac{\partial \tilde{f}}{\partial x_j} = \frac{\partial f}{\partial x_j} - \phi'(h)\frac{\partial h}{\partial x_j} = (1 - \phi'(h))\frac{\partial f}{\partial x_j} + \phi'(h)\frac{\partial g}{\partial x_j}.$$

Exactly as in the case of 2 players, $0 \leq \phi'(h) \leq 1$, and $\frac{\partial f}{\partial x_j}, \frac{\partial g}{\partial x_j}$ are both nonnegative. So, $\frac{\partial \tilde{f}}{\partial x_j} \geq 0$. For the second partial derivatives, we get

$$
\begin{aligned}
\frac{\partial^2 \tilde{f}}{\partial x_i \partial x_j} &= \frac{\partial^2 f}{\partial x_i \partial x_j} - \phi'(h)\frac{\partial^2 h}{\partial x_i \partial x_j} - \phi''(h)\frac{\partial h}{\partial x_i}\frac{\partial h}{\partial x_j} \\
&= (1 - \phi'(h))\frac{\partial^2 f}{\partial x_i \partial x_j} + \phi'(h)\frac{\partial^2 g}{\partial x_i \partial x_j} - \phi''(h)\frac{\partial h}{\partial x_i}\frac{\partial h}{\partial x_j}.
\end{aligned}
$$

The first two terms form a convex combination of non-positive values. To control the third term, we have $|\phi''(h)| \leq \alpha/h$. We also showed $\left|\frac{\partial h}{\partial x_i}\right| \leq k^3(1 - \bar{x})^{k/2-1}\sqrt{h(\mathbf{x})}$ (Claim 3). We can conclude that

$$\frac{\partial^2 \tilde{f}}{\partial x_i \partial x_j} \leq \left|\phi''(h)\frac{\partial h}{\partial x_i}\frac{\partial h}{\partial x_j}\right| \leq \alpha k^6(1 - \bar{x})^{k-2}.$$

We need to make the second partial derivatives non-positive. Since $\frac{\partial^2 g}{\partial x_i \partial x_j} = -\frac{k-1}{k}(1 - \bar{x})^{k-2}$, it is enough to add a suitable multiple of g to both functions: $\hat{f} = \tilde{f} + 2\alpha k^6 g$, $\hat{g} = (1 + 2\alpha k^6)g$. Then $\hat{f}, \hat{g}$ are smooth submodular.

Recall that we have $\alpha = 2/\ln \frac{1}{k\epsilon}$. For a given $\beta > 0$, we choose $\epsilon = \frac{1}{k}e^{-4k^6/\beta}$, so that $\beta = 2\alpha k^6$ and we increase g only by a factor of $1 + \beta$. We also get $\epsilon_2 = \sqrt{k\epsilon} \leq \beta$, and therefore $\hat{f}(\mathbf{x}) \geq \tilde{f}(\mathbf{x}) \geq f(\mathbf{x}) - \epsilon_2 \geq f(\mathbf{x}) - \beta$. Thus $\hat{f}$ and $\hat{g}$ satisfy the conditions of the lemma. $\qquad\square$

3 Value-Query Complexity of Subadditive Welfare Maximization

In this section, we construct an example showing that it is impossible to achieve an approximation factor asymptotically better than $\frac{1}{\sqrt{m}}$ for the subadditive welfare problem, using a polynomial number of value queries.

In fact, we prove our result for the more restricted class of fractionally subadditive valuation functions [12, 6] that is known to strictly contain all submodular valuation functions. A fractionally subadditive function is the pointwise maximum over a set of linear valuation functions.

Definition 3.1. *A linear valuation function (also known as additive) is a set function $a : 2^M \to \mathbb{R}_+$ that assigns a non-negative value to every singleton $\{j \in M\}$, and for all $S \subseteq M$ it holds that $a(S) = \sum_{j \in S} a(\{j\})$.*

Definition 3.2. *A fractionally subadditive function is a set function $f : S^M \to \mathbb{R}_+$, for which there is a finite set of linear valuation functions $A = \{a_1, ..., a_l\}$ such that $f(S) = \max_{a_i \in A} a_i(S)$, for every $S \subseteq M$.*

We prove the following theorem.

Theorem 3.1. *For any fixed $\epsilon > 0$, an $\frac{1}{m^{\frac{1}{2} - \epsilon}}$-approximation algorithm for fractionally subadditive valuation functions requires exponentially many value queries.*

Proof. We shall use probabilistic construction arguments to show that any algorithm that obtains an $\frac{1}{m^{\frac{1}{2} - \epsilon}}$-approximation to the social-welfare requires an exponential number of value queries.

Fix a small constant $\delta > 0$ (to be determined later). We shall construct a combinatorial auction with m items and $k = \sqrt{m}$ bidders. For every S let a_S be the linear valuation function that assigns a value of 1 to each item $j \in S$, and 0 to each item $j \notin S$. Let $\bar{a}$ be the additive valuation that assigns every item $j \in [m]$ a value of $\frac{1+\delta}{m^{\frac{1}{2} - \delta}}$.

Let $v_1, ..., v_k$ be an k-tuple of valuation functions defined as follows:

$$v_i = \max\{a_S\}_{|S| \leq (1+\delta)m^{2\delta}} \cup \{\bar{a}\}$$

Choose, uniformly at random, a partition of the items into $\sqrt{m}$ disjoint bundles of items $T_1, \ldots, T_k$ such that for each i $|T_i| = \sqrt{m}$. Let $v'_1, \ldots, v'_k$ be the k-tuple of valuation functions defined as follows:

$$v'_i = \max\{v_i, a_{T_i}\}$$

We shall now prove that for every player i, it takes an exponential number of value queries to distinguish between the case that i's valuation function is v_i and the case that i's valuation function is v'_i.

It is easy to see that the maximum social-welfare attainable if the valuation functions are $v_1 \ldots, v_k$ is $O(m^{\frac{1}{2} + 2\delta})$, while the optimal social-welfare if the valuation functions are $v'_1 \ldots, v'_k$ is m. Hence, the fact that it requires an exponential number of value-queries to distinguish between the valuation-functions profiles $v_1 \ldots, v_n$ and $v'_1 \ldots, v'_n$ implies that one cannot get an approximation ratio better than $O(m^{\frac{1}{2} - 2\delta})$ in less than an exponential number of value queries.

Consider a specific player i. Fix a bundle S of size smaller or equal to $m^{\frac{1}{2} + \delta}$. It holds that $v_i(S) = \max\{|S|, (1+\delta)m^{2\delta}\}$. v'_i might assign a value higher than v_i to S but only if

$$|S \cap T_i| > (1+\delta)m^{2\delta}.$$

Using standard probabilistic arguments, and relying on the Chernoff bounds, it can be shown that $Pr[|S \cap T_i| > (1+\delta)m^{2\delta}]$ is exponentially small.

Now, consider a bundle S of size greater than $m^{\frac{1}{2} + \delta}$. v_i will assign S the value of $(1+\delta)\frac{|S|}{m^{\frac{1}{2} - \delta}}$. v'_i might assign S a higher value, but only if

7

$$|S \cap T_i| > (1 + \delta)\frac{|S|}{m^{\frac{1}{2}-\delta}}.$$

Again, using standard probabilistic arguments it can be shown that $Pr[|S \cap T_i| > (1+\delta)\frac{|S|}{m^{\frac{1}{2}-\delta}}]$ is exponentially small.

We conclude that for every bundle S, only with exponentially small probability does one gather sufficient information to distinguish between the case that i's valuation is v_i and the case that it is v_i'. Hence, it might require an exponential number of value queries to distinguish between v_i and v_i'. This concludes the proof of the theorem. $\square$

Remark 3.2. *The proof of Theorem 3.1 can easily be made to hold for the case that all bidders have the same valuation functions.*

4 Value-Query Complexity of Superadditive Welfare Maximization

In this section, we construct an example showing that it is impossible to achieve an approximation factor asymptotically better than $\frac{\log m}{m}$ for the superadditive welfare problem, using a polynomial number of value queries.

The construction of the example will be done in two steps. First, we shall define a subclass of superadditive valuation functions we term *min-linear functions*. This is a superadditive analogue of fractionally subadditive functions [12, 6]. We shall then prove our lower bound for this more restricted class.

Definition 4.1. *A min-linear function is a set function $f : 2^M \to \mathbb{R}_+$ such that there is a finite set of linear valuation functions $A = \{a_1, \ldots, a_l\}$ such that for every $S \subseteq M$ $f(S) = \min_{a_i \in A} a_i(S)$.*

It is easy to show (and analogous to the proofs in [12, 11] that min-linear functions are contained in the class of superadditive valuation functions, and are a superclass of supermodular valuation functions. Simple examples demonstrate that these containments are strict.

Claim 4.1. *Any min-linear function is superadditive.*

Proof. Let f be a min-linear function. Let S and T be two disjoint subsets of items. By definition there are linear functions a_S, a_T, and $a_{S \cup T}$ in A for which the value of S, T, and $S \cup T$, is minimized. Therefore, it must hold that $a_S(S) \leq a_{S \cup T}(S)$ and $a_T(T) \leq a_{S \cup T}(T)$. Hence:

$$f(S \cup T) = a_{S \cup T}(S \cup T) = a_{S \cup T}(S) + a_{S \cup T}(T) \geq a_S(S) + a_T(T) = f(S) + f(T).$$

$\square$

Definition 4.2. *A set function f is supermodular iff for any two sets S and T, $f(S) + f(T) \leq f(S \cup T) + f(S \cap T)$.*

Claim 4.2. *Any supermodular valuation function is a min-linear function.*

Proof. Let f be a supermodular valuation function. Fix an order on the items, w.l.o.g., $1, \ldots, m$. For every set S we define a linear function a_S as follows: For every $j \in S$ $a_S(\{j\}) = f(\{1, \ldots, j\}) - f(\{1, \ldots, j-1\})$. For every $j \notin S$ $a_S(\{j\}) = \infty$, where ∞ represents a very high number (in particular $f(M) << \infty$). It is easy to see that $a_S(S) = f(S)$ for any S. We want to show that f

8

is min-linear and the finite set of linear functions is $A = \{a_T\}_{T \subseteq M}$. For this, we need to show that $a_S(S) = \min_{a_T \in A} a_T(S)$.

Observe, that for any S, T such that S is not contained in T it is impossible that the minimum for S is achieved by a_T (because for some $j \in S$ $a_T(\{j\}) = \infty$). So, we are left with the case that $S \subseteq T$. Here we exploit the well known fact that monotone supermodular functions have increasing marginal utilities. That is, if $U \subset V \subseteq M$, and j is in neither U nor V, then $f(U \cup \{j\}) - f(U) \leq f(V \cup \{j\}) - f(V)$. It is easy to see that this implies that for $S \subseteq T$, for any $j \in S$ $a_S(\{j\}) \leq a_T(\{j\})$. Hence, a_S is indeed the linear function for which the minimum is achieved (for bundle S). This implies that $f = \min A$. $\qquad\square$

We are now ready to prove the following theorem:

Theorem 4.1. *For any $\epsilon > 0$, a $\frac{\log^{1+\epsilon} m}{m}$-approximation algorithm for min-linear valuation functions requires a superpolynomial number of value queries.*

Proof. We use a probabilistic construction. Fix $\epsilon > 0$. We construct an auction with $|M| = m + \frac{m}{\log^{1+\epsilon} m}$ items and $|N| = k = \frac{m}{\log^{1+\epsilon} m}$. M consists of two disjoint sets M_1 and M_2, such that $|M_1| = m$ and $|M_2| = \frac{m}{\log^{1+\epsilon} m}$. Each bidder $i \in [k]$ is associated with a unique item d_i in M_2. We partition M_1, uniformly at random, into $k = \frac{m}{\log^{1+\epsilon} m}$ bundles $T_1, ..., T_k$ of equal size (i.e., of size $\log^{1+\epsilon} m$).

For each bidder i, let b_i be the linear valuation function that assigns a value of 1 to d_i and 0 to all other items. By a_S, we denote a linear function that assigns a value of 1 to each item in S and 0 to all other items.

Let $v_1, ..., v_k$ be the k-tuple of valuation functions defined as follows:

$$v_i = \min\{b_i\} \cup \{a_S : S \subseteq M_1, |S| = \frac{m}{2}\}.$$

This function has a very simple interpretation: It assigns a value of 1 to every set that contains d_i and more than $\frac{m}{2}$ items in M_1. It assigns 0 to all other bundles.

Let $v'_1, ..., v'_k$ be the k-tuple of valuation functions defined as follows:

$$v'_i = \min\{b_i\} \cup \{a_S : S \subseteq M_1, |S| = \frac{m}{2} \text{ \& } S \cap T_i \neq \emptyset\}.$$

This too has a simple interpretation: It assigns a value of 1 to every set that contains d_i and T_i, or d_i and more than $\frac{m}{2}$ items in M_1. It assigns 0 to all other sets.

We shall show that it might take a super-polynomial number of value queries to distinguish between the case that i's valuation function is v_i and the case that it is v'_i. It is easy to see that if the valuation functions of the bidders are $v_1, ..., v_k$ then the optimal social welfare is 1. On the other hand, if the valuation functions of the players are $v'_1, ..., v'_k$ then the optimal social welfare value is $\frac{m}{\log^{1+\epsilon} m}$ (assign every bidder i the bundle that contains d_i and T_i). Hence, it follows that achieving an approximation of $\frac{\log^{1+\epsilon} m}{m}$ requires a super-polynomial number of value queries (required to distinguish between $v_1, ..., v_k$ and $v'_1, ..., v'_k$).

Observe that, for each i, v_i and v'_i assign exactly the same value to all bundles, except for bundles that contain d_i and at most $\frac{m}{2}$ items in M_1. Also observe, that the difference is that v_i assigns a value of 0 to all these bundles, while v'_i assigns a value of 1 to such bundles that contain T_i (and 0 to all other such sets). What is the probability that a set of size at most $\frac{m}{2}$ in M_1 contains T_i? Let S be a bundle in M_1 of size at most $\frac{m}{2}$. Recall that T_i is uniformly distributed over all sets of size $\log^{1+\epsilon} m$. For every item in T_i, the probability that it is contained in S is at most $\frac{1}{2}$. Therefore,

$$Pr[T_i \subseteq S] \leq \frac{1}{2^{\log^{1+\epsilon} m}}$$

This implies that $2^{\log^{1+\epsilon} m}$, i.e., a super-polynomial number of value queries, may be required to distinguish between v_i and v_i'. The theorem follows.

$\square$

Remark 4.2. *The proof of Theorem 4.1 can easily be made to hold for the case that all bidders have the same valuation functions.*

5 Open Questions

We conclude by the following two open questions:

- For the case of submodular valuation functions, the only information-theoretic lower bound in the models of general and demand queries is $1 - O(1/m)$. There are indications that suggest the existence of a constant $(1 - \epsilon)$ lower bound (an NP-hardness result [3], and an integrality gap in [4]). Proving such a lower bound in any of these two models is a very interesting open question and seems to require non-trivial combinatorial constructions.

- We have presented tight lower bounds in the value query model for submodular and sub-additive valuation functions. There is still a gap between $\frac{\sqrt{\log m}}{m}$ and $\frac{\log m}{m}$ in the case of superadditive valuations. We have not considered the class of supermodular valuation functions, for which no information-theoretic lower bound is known in any of the models. Proving information-theoretic lower bounds for this class is an open problem.

References

[1] L. Blumrosen, and N. Nisan On the Computational Power of Iterative Auctions I: Demand Queries. In *Proceedings of the 6th ACM Conference on Electronic Commerce (EC)*, 2005.

[2] L. Blumrosen and N. Nisan. On the Computational Power of Iterative Auctions I: Ascending Auctions. In *Proceedings of the 6th ACM Conference on Electronic Commerce (EC)*, 2005.

[3] S. Dobzinski. Private communication.

[4] S. Dobzinski and M. Schapira. An Improved Approximation Algorithm for Combinatorial Auctions With Submodular Bidders In *Proceedings of the 18rd Annual ACM Symposium on Discrete Algorithms (SODA)*, 2006.

[5] S. Dobzinski, M. Schapira and N. Nisan. Approximation Algorithms for Combinatorial Auctions with Complement-Free Bidders In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing (STOC)*, 2005.

[6] U. Feige. On maximizing welfare when utility functions are subadditive. In *Proceedings of the 38rd Annual ACM Symposium on Theory of Computing (STOC)*, 2006.

[7] U. Feige and J. Vondrák. Approximation algorithms for combinatorial allocation problems: Improving the factor of $1 - 1/e$. In *Proc. of the 47th FOCS* (2006), 667–676.

[8] U. Feige, V. Mirrokni and J. Vondrák. Maximizing non-monotone submodular functions. In *Proceedings of 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2007.

[9] R. Holzman, N. Kfir-Dahav, D. Monderer and M. Tennenholtz. Bundling Equilibrium in Combinatorial Auctions. Games and Economic Behavior, 47, 104-123, 2004.

[10] S. Khot, R. Lipton, E. Markakis and A. Mehta. Inapproximability results for combinatorial auctions with submodular utility functions, in *Proceedings of WINE 2005*.

[11] B. Lehmann, D. Lehmann and N. Nisan. Combinatorial Auctions with Decreasing Marginal Utilities ACM EC 2001.

[12] N. Nisan. Bidding and allocation in combinatorial auctions. In Proceedings of the 2nd ACM Conference on Electronic Commerce (*EC*), Minneapolis, MN, 2000.

[13] N. Nisan and I. Segal. The Communication Requirements of Efficient Allocations and Supporting Prices. Journal of Economic Theory, 129(1), pp. 192-224, , July 2006

[14] N. Nisan and I. Segal. Exponential Communication Inefficiency of Demand Queries. Theoretical Aspects of Rationality and Knowledge X, June 2005

[15] G. Nemhauser, L. Wolsey and M. Fisher. An analysis of the approximations for maximizing submodular set functions. Mathematical Programming, 14, pp. 265-294, 1978.

[16] M. Fisher, G. Nemhauser and L. Wolsey. An analysis of approximations for maximizing submodular set functions II. Math. Programming Study 8 (1978), 73–87.

[17] J. Vondrák. Optimal approximation for the Submodular Welfare Problem in the value oracle model, *submitted*, 2007.